SEQUENCE LISTING

<110> van Rooijen, Gijs
Keon, Richard Glenn
Boothe, Joseph
Shen, Yin

<120> Commercial Production of Chymosin in Plants

<130> 9369-153

<140> 09/643,755
<141> 2000-08-23

<160> 4

<170> PatentIn Ver. 2.0

<210> 1
<211> 1173
<212> DNA
<213> Bovine

<220>
<221> CDS
<222> (1)..(1173)

<400> 1

atg	aac	ttc	ctt	aag	tct	ttc	cct	ttc	tac	gct	ttc	ctt	tgt	ttc	ggt	48
Met	Asn	Phe	Leu	Lys	Ser	Phe	Pro	Phe	Tyr	Ala	Phe	Leu	Cys	Phe	Gly	
1				5					10					15		

caa	tac	ttc	gtt	gct	gtt	act	cac	gct	gct	gag	atc	acc	cgc	att	cct	96
Gln	Tyr	Phe	Val	Ala	Val	Thr	His	Ala	Ala	Glu	Ile	Thr	Arg	Ile	Pro	
			20					25					30			

ctc	tac	aaa	ggt	aag	tct	ctc	cgt	aag	gcg	ctg	aag	gaa	cat	gga	ctt	144
Leu	Tyr	Lys	Gly	Lys	Ser	Leu	Arg	Lys	Ala	Leu	Lys	Glu	His	Gly	Leu	
		35					40					45				

cta	gaa	gac	ttc	ttg	cag	aaa	caa	cag	tat	ggc	atc	agc	agc	aag	tac	192
Leu	Glu	Asp	Phe	Leu	Gln	Lys	Gln	Gln	Tyr	Gly	Ile	Ser	Ser	Lys	Tyr	
	50					55					60					

tcc	ggc	ttc	ggt	gaa	gtt	gct	agc	gtg	cca	ctt	acc	aac	tac	ctt	gat	240
Ser	Gly	Phe	Gly	Glu	Val	Ala	Ser	Val	Pro	Leu	Thr	Asn	Tyr	Leu	Asp	
65					70					75					80	

agt	caa	tac	ttt	ggg	aag	atc	tac	ctc	gga	acc	ccg	cct	caa	gag	ttc	288
Ser	Gln	Tyr	Phe	Gly	Lys	Ile	Tyr	Leu	Gly	Thr	Pro	Pro	Gln	Glu	Phe	
				85					90					95		

acc	gtt	ctc	ttt	gat	act	ggt	tcc	tct	gac	ttc	tgg	gtt	ccc	tct	atc	336
Thr	Val	Leu	Phe	Asp	Thr	Gly	Ser	Ser	Asp	Phe	Trp	Val	Pro	Ser	Ile	
			100					105					110			

tac	tgc	aag	agc	aat	gcc	tgc	aag	aac	cac	caa	aga	ttc	gat	ccg	aga	384
Tyr	Cys	Lys	Ser	Asn	Ala	Cys	Lys	Asn	His	Gln	Arg	Phe	Asp	Pro	Arg	
		115					120					125				

aag	tcg	tcc	acc	ttc	cag	aac	tta	ggc	aaa	ccc	ttg	tct	ata	cac	tac	432
Lys	Ser	Ser	Thr	Phe	Gln	Asn	Leu	Gly	Lys	Pro	Leu	Ser	Ile	His	Tyr	
	130					135					140					

| ggt | aca | ggt | agc | atg | caa | gga | atc | tta | ggc | tat | gat | acc | gtc | act | gtc | 480 |

```
Gly Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val
145             150             155             160

tcc aac att gtg gac att caa cag aca gta gga ctt agc acc caa gaa    528
Ser Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu
            165             170             175

cca ggt gat gtc ttc acc tat gca gaa ttc gat ggc atc ctt ggt atg    576
Pro Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met
            180             185             190

gca tac cca tcg ctc gcg tca gag tac tcg ata cct gtg ttt gac aac    624
Ala Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn
            195             200             205

atg atg aac cga cac cta gta gct caa gac ttg ttc tcg gtt tac atg    672
Met Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met
210             215             220

gac agg aat ggc cag gag agc atg ctc acg ctt gga gct att gat cca    720
Asp Arg Asn Gly Gln Glu Ser Met Leu Thr Leu Gly Ala Ile Asp Pro
225             230             235             240

tcc tac tac aca gga tct ctt cac tgg gtt cca gtc act gtg cag cag    768
Ser Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln
            245             250             255

tac tgg caa ttc act gtg gac agt gtc acc atc agc ggt gtg gtt gtt    816
Tyr Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val
            260             265             270

gca tgt gaa ggt gga tgt caa gct atc ttg gat acc ggt acg tcc aag    864
Ala Cys Glu Gly Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys
            275             280             285

ctg gtc gga cct agc agc gac att ctc aac att cag caa gct att gga    912
Leu Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly
    290             295             300

gcc aca cag aac cag tac ggt gag ttt gac ata gat tgc gac aac ctt    960
Ala Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu
305             310             315             320

agc tac atg cct aca gtt gtc ttt gag atc aac ggc aag atg tac cca   1008
Ser Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro
            325             330             335

ctg acc ccc tcc gcc tat acc agc cag gat caa ggg ttc tgc acc agt   1056
Leu Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser
            340             345             350

gga ttc cag agt gag aac cat tcc cag aaa tgg atc ttg gga gat gtg   1104
Gly Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val
            355             360             365

ttc att cgt gag tac tac agc gtc ttt gac agg gcc aac aac ctc gtt   1152
Phe Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val
    370             375             380

ggg cta gct aaa gca atc tga                                       1173
Gly Leu Ala Lys Ala Ile
385             390
```

```
<210> 2
<211> 390
```

```
<212> PRT
<213> Bovine

<400> 2
Met Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly
 1               5                  10                  15

Gln Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro
             20                  25                  30

Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu
         35                  40                  45

Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr
     50                  55                  60

Ser Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp
 65                  70                  75                  80

Ser Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe
                 85                  90                  95

Thr Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile
             100                 105                 110

Tyr Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg
         115                 120                 125

Lys Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr
     130                 135                 140

Gly Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val
145                 150                 155                 160

Ser Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu
                 165                 170                 175

Pro Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met
             180                 185                 190

Ala Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn
         195                 200                 205

Met Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met
     210                 215                 220

Asp Arg Asn Gly Gln Glu Ser Met Leu Thr Leu Gly Ala Ile Asp Pro
225                 230                 235                 240

Ser Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln
                 245                 250                 255

Tyr Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val
             260                 265                 270

Ala Cys Glu Gly Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys
         275                 280                 285

Leu Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly
     290                 295                 300

Ala Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu
305                 310                 315                 320

Ser Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro
             325                 330                 335
```

Leu Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser
       340   345   350

Gly Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val
       355   360   365

Phe Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val
    370   375   380

Gly Leu Ala Lys Ala Ile
385   390

<210> 3
<211> 3957
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1554)..(2726)

<220>
<223> Description of Artificial Sequence: Phaseolin promoter- pre-pro-
                                   chymosin-phaseolin terminator

<400> 3
ctgcaggaat tcattgtact cccagtatca ttatagtgaa agttttggct ctctcgccgg 60

tggtttttta cctctatttta aaggggtttt ccacctaaaa attctggtat cattctcact 120

ttacttgtta ctttaatttc tcataatctt tggttgaaat tatcacgctt ccgcacacga 180

tatccctaca aatttattat ttgttaaaca ttttcaaacc gcataaaatt ttatgaagtc 240

ccgtctatct ttaatgtagt ctaacatttt catattgaaa tatataattt acttaatttt 300

agcgttggta gaaagcataa agatttattc ttattcttct tcatataaat gtttaatata 360

caatataaac aaattcttta ccttaagaag gatttcccat tttatatttt aaaaatatat 420

ttatcaaata tttttcaacc acgtaaatct cataataata agttgtttca aaagtaataa 480

aatttaactc cataattttt ttattcgact gatcttaaag caacacccag tgacacaact 540

agccatttt ttctttgaat aaaaaaatcc aattatcatt gtattttttt tatacaatga 600

aaatttcacc aaacaatcat ttgtggtatt tctgaagcaa gtcatgttat gcaaaattct 660

ataattccca tttgacacta cggaagtaac tgaagatctg cttttacatg cgagacacat 720

cttctaaagt aattttaata atagttacta tattcaagat ttcatatatc aaatactcaa 780

tattacttct aaaaaattaa ttagatataa ttaaaatatt actttttaa ttttaagttt 840

aattgttgaa tttgtgacta ttgatttatt attctactat gtttaaattg ttttatagat 900

agtttaaagt aaatataagt aatgtagtag agtgttagag tgttaccccta aaccataaac 960

tataacattt atggtggact aattttcata tatttcttat tgcttttacc ttttcttggt 1020

atgtaagtcc gtaactagaa ttacagtggg ttgccatggc actctgtggt ctttttggttc 1080

atgcatgggt cttgcgcaag aaaaagacaa agaacaaaga aaaaagacaa aacagagaga 1140

caaaacgcaa tcacacaacc aactcaaatt agtcactggc tgatcaagat cgccgcgtcc 1200

atgtatgtct aaatgccatg caaagcaaca cgtgcttaac atgcacttta aatggctcac 1260

ccatctcaac ccacacacaa acacattgcc tttttcttca tcatcaccac aaccacctgt 1320

atatattcat tctcttccgc cacctcaatt tcttcacttc aacacacgtc aacctgcata 1380

tgcgtgtcat cccatgccca aatctccatg catgttccaa ccaccttctc tcttatataa 1440

tacctataaa tacctctaat atcactcact tctttcatca tccatccatc cagagtacta 1500

ctactctact actataatac cccaacccaa ctcatattca atactactct act atg   1556
                                                           Met
                                                           1

aac ttc ctt aag tct ttc cct ttc tac gct ttc ctt tgt ttc ggt caa   1604
Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly Gln
            5   10   15

tac ttc gtt gct gtt act cac gct gct gag atc acc cgc att cct ctc   1652
Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro Leu
        20   25   30

tac aaa ggt aag tct ctc cgt aag gcg ctg aag gaa cat gga ctt cta   1700
Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu Leu
    35   40   45

gaa gac ttc ttg cag aaa caa cag tat ggc atc agc agc aag tac tcc   1748
Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr Ser
50   55   60   65

ggc ttc ggt gaa gtt gct agc gtg cca ctt acc aac tac ctt gat agt   1796
Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp Ser
                70   75   80

caa tac ttt ggg aag atc tac ctc gga acc ccg cct caa gag ttc acc   1844
Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe Thr
            85   90   95

gtt ctc ttt gat act ggt tcc tct gac ttc tgg gtt ccc tct atc tac   1892
Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile Tyr
        100   105   110

tgc aag agc aat gcc tgc aag aac cac caa aga ttc gat ccg aga aag   1940
Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg Lys
    115   120   125

tcg tcc acc ttc cag aac tta ggc aaa ccc ttg tct ata cac tac ggt   1988
Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr Gly
130   135   140   145

aca ggt agc atg caa gga atc tta ggc tat gat acc gtc act gtc tcc   2036
Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val Ser
                150   155   160

aac att gtg gac att caa cag aca gta gga ctt agc acc caa gaa cca   2084
Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu Pro
                165   170   175

ggt gat gtc ttc acc tat gca gaa ttc gat ggc atc ctt ggt atg gca   2132
Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met Ala
            180   185   190

tac cca tcg ctc gcg tca gag tac tcg ata cct gtg ttt gac aac atg   2180
Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn Met
    195   200   205

```
atg aac cga cac cta gta gct caa gac ttg ttc tcg gtt tac atg gac     2228
Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met Asp
210                 215                 220                 225

agg aat ggc cag gag agc atg ctc acg ctt gga gct att gat cca tcc     2276
Arg Asn Gly Gln Glu Ser Met Leu Thr Leu Gly Ala Ile Asp Pro Ser
                    230                 235                 240

tac tac aca gga tct ctt cac tgg gtt cca gtc act gtg cag cag tac     2324
Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln Tyr
                245                 250                 255

tgg caa ttc act gtg gac agt gtc acc atc agc ggt gtg gtt gtt gca     2372
Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val Ala
            260                 265                 270

tgt gaa ggt gga tgt caa gct atc ttg gat acc ggt acg tcc aag ctg     2420
Cys Glu Gly Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys Leu
        275                 280                 285

gtc gga cct agc agc gac att ctc aac att cag caa gct att gga gcc     2468
Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly Ala
290                 295                 300                 305

aca cag aac cag tac ggt gag ttt gac ata gat tgc gac aac ctt agc     2516
Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu Ser
                310                 315                 320

tac atg cct aca gtt gtc ttt gag atc aac ggc aag atg tac cca ctg     2564
Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro Leu
                325                 330                 335

acc ccc tcc gcc tat acc agc cag gat caa ggg ttc tgc acc agt gga     2612
Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser Gly
            340                 345                 350

ttc cag agt gag aac cat tcc cag aaa tgg atc ttg gga gat gtg ttc     2660
Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val Phe
        355                 360                 365

att cgt gag tac tac agc gtc ttt gac agg gcc aac aac ctc gtt ggg     2708
Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val Gly
370                 375                 380                 385

cta gct aaa gca atc tga agcttaataa gtatgaacta aaatgcatgt           2756
Leu Ala Lys Ala Ile
                390
```

```
aggtgtaaga gctcatggag agcatggaat attgtatccg accatgtaac agtataataa 2816

ctgagctcca tctcacttct tctatgaata aacaaaggat gttatgatat attaacactc 2876

tatctatgca ccttattgtt ctatgataaa tttcctctta ttattataaa tcatctgaat 2936

cgtgacggct tatggaatgc ttcaaatagt acaaaaacaa atgtgtacta taagactttc 2996

taaacaattc taactttagc attgtgaacg agacataagt gttaagaaga cataacaatt 3056

ataatggaag aagtttgtct ccatttatat attatatatt acccacttat gtattatatt 3116

aggatgttaa ggagacataa caattataaa gagagaagtt tgtatccatt tatatattat 3176

atactaccca tttatatatt atacttatcc acttatttaa tgtctttata aggtttgatc 3236

catgatattt ctaatatttt agttgatatg tatatgaaag ggtactattt gaactctctt 3296
```

```
actctgtata aaggttggat catccttaaa gtgggtctat ttaattttat tgcttcttac 3356

agataaaaaa aaaattatga gttggtttga taaaatattg aaggatttaa aataataata 3416

aataataaat aacatataat atatgtatat aaatttatta taatataaca tttatctata 3476

aaaaagtaaa tattgtcata aatctataca atcgtttagc cttgctggac gactctcaat 3536

tatttaaacg agagtaaaca tatttgactt tttggttatt taacaaatta ttatttaaca 3596

ctatatgaaa tttttttttt ttatcggcaa ggaaataaaa ttaaattagg agggacaatg 3656

gtgtgtccca atccttatac aaccaacttc cacaggaagg tcaggtcggg gacaacaaaa 3716

aaacaggcaa gggaaatttt ttaatttggg ttgtcttgtt tgctgcataa tttatgcagt 3776

aaaacactac acataaccct tttagcagta gagcaatggt tgaccgtgtg cttagcttct 3836

tttattttat ttttttatca gcaaagaata aataaaataa aatgagacac ttcagggatg 3896

tttcaaccct tatacaaaac cccaaaaaca agttcctag cacccctacca actaaggtac 3956

c                                                                 3957
```

<210> 4
<211> 390
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Phaseolin promoter- pre-pro-
                                          chymosin-phaseolin terminator

<400> 4

```
Met Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly
 1               5                  10                  15

Gln Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro
            20                  25                  30

Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu
        35                  40                  45

Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr
    50                  55                  60

Ser Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp
65                  70                  75                  80

Ser Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe
                85                  90                  95

Thr Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile
            100                 105                 110

Tyr Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg
        115                 120                 125

Lys Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr
    130                 135                 140

Gly Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val
145                 150                 155                 160

Ser Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu
```

```
                    165                    170                    175
Pro Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met
            180                185                190

Ala Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn
            195                200                205

Met Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met
    210                215                220

Asp Arg Asn Gly Gln Glu Ser Met Leu Thr Leu Gly Ala Ile Asp Pro
225                230                235                240

Ser Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln
            245                250                255

Tyr Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val
            260                265                270

Ala Cys Glu Gly Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys
    275                280                285

Leu Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly
    290                295                300

Ala Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu
305                310                315                320

Ser Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro
            325                330                335

Leu Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser
            340                345                350

Gly Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val
            355                360                365

Phe Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val
    370                375                380

Gly Leu Ala Lys Ala Ile
385                390
```